SEQUENCE LISTING

5	(1) GE	NERAL INFORMATION:
	(i) APPLICANT: Gorski, David H. Walsh, Kenneth
10	i)	.i) TITLE OF INVENTION: Growth Arrest Homeobox Gene
1.0	(ii	i) NUMBER OF SEQUENCES: 4
15		(A) ADDRESSEE: Calfee, Halter, and Griswold (B) STREET: 800 Superior Avenue (C) CITY: Cleveland (D) STATE: Ohio (E) COUNTRY: U.S.A. (F) ZIP: 44114-2688
20 25 5	,	(v) COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
30	· .	(A) APPLICATION DATA: (B) FILING DATE: (C) CLASSIFICATION:
35	(vi	Li) ATTORNEY/AGENT INFORMATION: (A) NAME: Golrick, Mary E. (B) REGISTRATION NUMBER: 34829 (C) REFERENCE/DOCKET NUMBER: 22311/00114
40	(ix) TELECOMMUNICATION INFORMATION: (A) TELEPHONE: (216) 622-8200 (B) TELEFAX: (216) 241-0816 (C) TELEX: 980499
	(2) I	NFORMATION FOR SEQ ID NO:1:
45		(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2244 base pairs
•		(B) TYPE: nucleic acid (C) STRANDEDNESS: both (D) TOPOLOGY: linear
50	,	ii) MOLECULE TYPE: cDNA
1/	•	ii) HYPOTHETICAL: NO
55	•	iv) ANTI-SENSE: NO
	`	•
60		ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 1971108
		xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
65	GTCAZ 60	GTGTT TATACGTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT

and field at the first manifestable and the complete of the control of the state of the control of the control

		CTGC 120	TCAA	AC C	CCGCC	CGGC	T T	TAC	ATTAC	GA(FTGA	GTGG	GGG	AGAG'	rcc	TAGG.	ATTTCT
<u>5</u>		AGTG 180	AAAA	AGT C	ACAC	CGCI	T GO	FTGG1	ACTT	r GGG	FACC	TTCG	TGA	AGTC'	TTC	TGCT	TGGAAG
		CTGA 229	GACI	TG (CATGO												
10						Me	t Gl 1	u Hi	s Pr	o Le	u Ph 5	e Gl	у Су	's Le	_	g Se .0	r
	٠.	CCC 277	CAC	GCC	ACA	GCG	CAA	GGC	TTG	CAC	CCC	TTC	TCG	CAG	TCT	TCT	CTG
15		Pro	His	Ala	Thr 15	Ala	Gln	Gly	Leu	His 20	Pro	Phe	Ser	Gln	Ser 25	Ser	Leu
		GCC 325	CTC	CAT	GGA	AGA	TCT	GAC	CAC	ATG	TCC	TAC	ccc	GAA	CTC	TCC	ACA
20		Ala	Leu	His 30	GLY	Arg	Ser	Asp	His 35	Met	Ser	Tyr	Pro	Glu 40	Leu	Ser	Thr
2 5		TCT 373	TCC	TCG	TCT	TGC	ATA	ATC	GCG	GGA	TAC	CCC	AAT	GAG	GAG	GGC	ATG
2 5		Ser	Ser 45	Ser	Ser	Суз	Ile	Ile 50	Ala	Gly	Tyr	Pro	Asn 55	Glu	Glu	Gly	Met
		TTT 421	GCC	AGC	CAG	CAT	CAC	AGG	GGG	CAC	CAC	CAC	CAC	CAC	CAC	CAC	CAC
30	,	Phe 60	Ala	Ser	Gln	His	His 65	Arg	Gly	His	His	His 70	His	His	His	His	His 75
2		CAT 469	CAC	CAC	CAC	CAG	CAG	CAG	CAG	CAC	CAG	GCT	CTG	CAA	AGC	AAC	TGG
35		His	His	His	His	Gln 80	Gln	Gln	Gln	His	Gln 85	Ala	Leu	Gln	Ser	Asn 90	Trp
		CAC 517	CTC	CCG	CAG	ATG	TCC	TCC	CCG	CCA	AGC	GCG	GCC	CGG	CAC	: AGC	CTT
40		His	Leu	Pro	Gln 95	Met	Ser	Ser	Pro	Pro 100	Ser	Ala [.]	Ala	Arg	His 105	Ser	Leu
		TGC 565	CTG	CAG	CCT	GAT	TCC	GGA	GGG	CCC	CCG	GAG	CTG	GGG	AGC	: AGC	CCT
45		Cys	Leu	Gln 110	Pro	Asp	Ser	_	Gly 115	Pro	Pro	Glu	Leu	Gly 120	Ser	Ser	Pro
		CCG 613	GTC	CTG	TGC	TCC	AAC	TCT	TCT	AGC	CTG	GGC	TCC	AGC	ACC	ccé	ACC
50		Pro	Val 125	Leu	Cys	Ser	Asn	Ser 130	Ser	Ser	Leu	Gly	Ser 135	Ser	Thr	Pro	Thr
		GGA 661	GCC	GCG	TĢC	GCA	CCA	AGG	GAT	TAT	GGC	CGT	CAA	GC6	CT	TCF	CCC
55			Ala	Ala	СЛа	Ala	Pro 145	Arg	Asp	Tyr	Gly	Arg 150	Gln	Ala	Leu	Ser	Pro 155
		GCA 709	GAA	GTG	GAG	AAG	AGA	AGT	GGC	AGC	AAA	A AGA	AAA	A AGO	GA(C AG	TCA
60			Glu	Val	Glu	Lys 160	Arg	Ser	Gly	Ser	Lys 165	Arg	Lys	Ser	Asp	Ser 170	Ser
		GAT 757	TCC	CAG	GAA	GGA	AAT	TAC	AAG	TCA	GAA	A GTG	AAC	C AGO	C AA	A CC	r agg
65			Ser	Gln	Glu 175	Gly	Asn	Tyr	Lys	Ser 180	Glu	Val	Asn	Ser	Lys 185	Pro	Arg

AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA

SEQUENCE LISTING

e	(1) GENERAL INFORMATION:
5	(i) APPLICANT: Gorski, David H. Walsh, Kenneth
10	(ii) TITLE OF INVENTION: Growth Arrest Homeobox Gene
	(iii) NUMBER OF SEQUENCES: 4
15	<pre>(iv) CORRESPONDENCE ADDRESS: (A) ADDRESSEE: Calfee, Halter, and Griswold (B) STREET: 800 Superior Avenue (C) CITY: Cleveland (D) STATE: Ohio (E) COUNTRY: U.S.A. (F) ZIP: 44114-2688</pre>
20	(r) hr: 44114-2000
25	 (v) COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
30	<pre>(vi) CURRENT APPLICATION DATA: (A) APPLICATION NUMBER: (B) FILING DATE: (C) CLASSIFICATION:</pre>
	<pre>(viii) ATTORNEY/AGENT INFORMATION: (A) NAME: Golrick, Mary E. (B) REGISTRATION NUMBER: 34829 (C) REFERENCE/DOCKET NUMBER: 22311/00114</pre>
35 40	(ix) TELECOMMUNICATION INFORMATION: (A) TELEPHONE: (216) 622-8200 (B) TELEFAX: (216) 241-0816 (C) TELEX: 980499
	(2) INFORMATION FOR SEQ ID NO:1:
45	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2244 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: both (D) TOPOLOGY: linear
50	(ii) MOLECULE TYPE: cDNA
i j	(iii) HYPOTHETICAL: NO
55	(iv) ANTI-SENSE: NO
60	(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 1971108
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
65	GTCAAGTGTT TATACGTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT

~~~	 100	
	 7 1 1 1 1	1 1 1
44-	 <i>1</i> u u	114

		CTGCTCAAAC CCGCGCGCT TTTACATTAG GAGTGAGTGG GGGAGAGTCC TAGGATTTCT
5	:	AGTGAAAAGT GACAGCGCTT GGTGGACTTT GGGACCTTCG TGAAGTCTTC TGCTTGGAAG 180
,		CTGAGACTTG CATGCC ATG GAA CAC CCC CTC TTT GGC TGC CTG CGC AGC
10		Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser 1 5 10
•		CCC CAC GCC ACA GCG CAA GGC TTG CAC CCC TTC TCG CAG TCT TCT CTG
15		Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu  15 20 25
		GCC CTC CAT GGA AGA TCT GAC CAC ATG TCC TAC CCC GAA CTC TCC ACA
20		Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr 30 35 40
		TCT TCC TCG TCT TGC ATA ATC GCG GGA TAC CCC AAT GAG GAG GGC ATG
25		Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met 45 50 55
<b>25</b>		TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC CAC CAC CAC
3.Q		Phe Ala Ser Gln His His Arg Gly His His His His His His His 60 65 70 75
my m	•	CAT CAC CAC CAG CAG CAG CAG CAC CAG GCT CTG CAA AGC AAC TGG
3 <b>.5</b>		His His His His Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp 80 85 90
Hand John Mary		CAC CTC CCG CAG ATG TCC TCC CCG CCA AGC GCG GCC CGG CAC AGC CTT
40		His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu 95 100 105
# .		TGC CTG CAG CCT GAT TCC GGA GGG CCC CCG GAG CTG GGG AGC AGC CCT
45		Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro 110 115 120
		CCG GTC CTG TGC TCC AAC TCT TCT AGC CTG GGC TCC AGC ACC CCG ACC 613
50·		Pro Val Leu Cys Ser Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr 125 130 135
•		GGA GCC GCG TGC GCA CCA AGG GAT TAT GGC CGT CAA GCG CTG TCA CCC 661
55		Gly Ala Ala Cys Ala Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro 140 145 150 155
		GCA GAA GTG GAG AAG AGA AGT GGC AGC AAA AGA AAA AGC GAC AGT TCA
60		Ala Glu Val Glu Lys Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser 160 165 170
		GAT TCC CAG GAA GGA AAT TAC AAG TCA GAA GTG AAC AGC AAA CCT AGG 757
65		Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg 175 180 185

AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA

	805
	Arg Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala 190 195 200
5	GAG TTC GCC CAT CAT AAC TAT CTG ACC AGA CTG AGA AGA TAT GAG ATA 853
•	Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile 205 210 215
10	GCG GTG AAC CTA GAC CTC ACT GAA AGA CAG GTG AAA GTG TGG TTC CAG
	Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln 220 235 .
15	AAC AGG AGA ATG AAG TGG AAG CGG GTC AAG GGG GGA CAA CAA GGA GCT 949
	Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala 240 245 250
20	GCA GCC CGA GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA ACA CTT CTT 997
	Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu 255 260 265
	CCA TCA GAG CTG TCA GGA ATT GGT GCA GCC ACC CTC CAG CAG ACA GGG
	Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly 270 275 280
<b>36</b>	GAC TCA CTA GCA AAT GAC GAC AGT CGC GAT AGT GAC CAC AGC TCT GAG
	Asp Ser Leu Ala Asn Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu 285 290 295
, marin	dia no mano dia non-
<b>35</b>	CAC GCA CAC TTA TGATACATAC AGAGACCAGC TCCGTTCTCA GGAAAGCACC 1145
	CAC GCA CAC TTA TGATACATAC AGAGACCAGC TCCGTTCTCA GGAAAGCACC 1145 His Ala His Leu 300
Maria Maria	His Ala His Leu
40	His Ala His Leu 300  ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG
Start Start	His Ala His Leu 300  ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205  CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC
40	His Ala His Leu 300  ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205  CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265  ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA
45	His Ala His Leu 300  ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205  CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265  ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA 1325  AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA
45	His Ala His Leu 300  ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205  CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265  ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA 1325  AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA 1385  AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATTT TGACTCAAGT TGTCTCCAGA
45 50 55	His Ala His Leu  300  ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG  1205  CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC  1265  ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA  1325  AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA  1385  AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATT TGACTCAAGT TGTCTCCAGA  1445  GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACACC
45	HIS ALA HIS Leu  300  ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG  1205  CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC  1265  ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA  1325  AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA  1385  AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATTT TGACTCAAGT TGTCTCCAGA  1445  GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC  TGGGCAGGTA TTTGCTTTTG CTTGCACTGA AACTTAAACT GCTATCAAGT TAACCCATGA
45 50 55	His Ala His Leu  300  ATTGTGATGG CAAATCTCAC CCAAACATCG TTTACATGGC AGATGACTGT GGCAGTGTTG 1205  CTTAATATAA TTAAACGCAG GCATCTCAAG TCTGTTTCTC ATGATTGATA GAAGGTTTAC 1265  ACTAAGTGCC TCTTATTGAA GATGCTTCCA CAGTGAAATT GGAGAAAGTG AACATATCTA 1325  AATATACTTG TTCCTTATAT GACAGAGAGG GAGATGAATG TTTGCTTTGG CTTGCACTGA 1385  AAATTAAATT GCTACCAAGA GCAAACTCGG TAAGACATT TGACTCAAGT TGTCTCCAGA 1445  GTGAAGATGT TATAGAAATG CTTTGAACAT TCCAGTTGTA CCAGGTCATG TGTGTGACAC 1505  TGGGCAGGTA TTTGCTTTTG CTTGCACTGA AACTTAAACT GCTATCAAGT TAACCCATGA  AAATAGTTTAT CTTGAACAGC CACAGTGCCT GAAATCACCA AGTGGATATA AAATGAACTG

•	~	A	
1	•	4	-

TTTATACATT TTTTTGTCAG TCAGAGACAT CAGTATGTGC TTACTTGTGT TCAAGTAGAG 1805

GAAATGCAGT AGAGTCTGAT AGGACATATT CTTGGTACCA CAGACAAAAC AAATCTTCTG

TTGCATTGAC TATCAACTGC TGCAGATACA TTAGAGAACA CACCTAGCCC CCCTCCAGCC 1925

TCCCTCTGTT ATCGCTCGAA GACATTAGCG TCATAGGCAA GTAGTTACCT TGCCAAATGA

GTCTTGTGTG GCAGATGTCT GATTTTGTAT CTTTAAACTG TTAATGGTAT GTGTCTGCTT 2045

CAGTTAACAG GGAAAAAGAT TTCTTCCTCA TTGTTTATGA TACAAAACCC AAGTGCCAAA 2105

CAAAGCTAGT TCTTCAAGGG ATAGATGAGA AACTGAATGT CTGACAAGTA GACTCAGCGA 2165

AAATACATTA TTTTTCAGAG GCTGTGTATT CATGCAGTAC AAGTCCTTGT ATTTTGTAAA 2225

AAAAAAGTT AAATAAATG 2244

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 303 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

#### (ii) MOLECULE TYPE: protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser Pro His Ala Thr Ala 1 5 .10 15

Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu Ala Leu His Gly Arg 20 25 30

Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr Ser Ser Ser Cys 35 40 45

Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met Phe Ala Ser Gln His 50 55 60

Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp His Leu Pro Gln Met 85 90 95

Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu Cys Leu Gln Pro Asp 100 105 110

Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro Pro Val Leu Cys Ser 115 120 125

Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr Gly Ala Ala Cys Ala 130 135 140

30 35 3

45

50.

55

60

65

5

10

15

20

22311	/00114
-------	--------

	Pro 145	Arg	Asp	Tyr	Gly	Arg 150	Gln	Ala	Leu	Ser	Pro 155	Ala	Glu	Val	Glu	Lys 160
<b>5</b>	Arg	Ser	Gly	Ser	Lys 165	Arg	Lys	Ser	Asp	Ser 170	Ser	Asp	Ser	Gln	Glu 175	Gly
	Asn	Tyr	Lys	Ser 180	Glu	Val	Asn	Ser	Lys 185	Pro	Arg	Arg	Glu	Arg 190	Thr	Ala
10	Phe	Thr	Lys 195	Glu	Gln	Ile	Arg	Glu 200	Leu	Glu	Ala	Glu	Phe 205	Ala	His	His
15	Asn	Tyr 210	Leu	Thr	Arg	Leu	Arg 215	Arg	Tyr	Glu	Ile	Ala 220	Val	Asn	Leu	Asp
. *	Leu 225	Thr	Glu	Arg	Gln	Val 230	Lys	Val	Trp	Phe	Gln 235	Asn	Arg	Arg	Met	Lys 240
20	Trp	ŗàa	Arg	Val	Lys 245	Gly	Gly	Gln	Gln	Gly 250	Ala	Ala	Ala	Arg	Glu 255	Lys
	Glu	Leu	Val	Asn 260	Val ·	Lys	Lys	Gly	Thr 265		Leu	Pro	Ser	Glu 270	Leu	Ser
25 40 5	Gly	Ile	Gly 275	Ala	Ala	Thr	Leu	Gln 280	Gln	Thr	Gly	Asp	Ser 285	Leu	Ala	Asn
ីា 3 <b>0</b>	Asp	Asp 290	Ser	Arg	Asp	Ser	Asp 295	His	Ser	Ser	Glu	His 300	Ala	His	Leu	٠
	(2)	INFO	RMAI	TION	FOR	SEQ	ID N	10:3:							•	٠.
Bertal Walls and Bridge and Bridg		(i)	(F (E	A) LE 3) TY C) ST	ength (PE : (Rani	i: 94 nucl	CTERI 11 ba Leic ESS: line	acio both	pairs i	3	•		•			
40		(ii)	MOI	LECUI	LE TY	PE:	CDNA	4								
	•	(iii)	HYI	POTHE	TIC	AL: 1	OF									
45		(iv)	ANT	ri—se	ense :	NO.										
50		(ix)	( <b>F</b>	ATURE A) NF B) LO	AME/I		CDS 33.	.941		٠						
		(xi)	SEÇ	QUENC	CE DI	ESCR	IPTIC	ON: S	SEQ I	ED NO	3:					
55	GTC: 53	rtct?	ACC 1	rggaž	ACCC	GA AZ	ACTTO	CATO	G CT						TTT Phe	
										1				5		_
60	101	CTG Leu														
			10					15					20			
65	149	TCC														
	Gln	Ser 25	Ser	Leu	Ala	Leu	His 30	Gly	Arg	Ser	Asp	His 35	Met	Ser	Tyr	Pro

	GAG	CTC	TCT	ACT	TCT	TCC	TCA	TCT	TGC	ATA	ATC	GCG	GGA	TAC	ccc	AAC
	197 Glu 40	Leu	Ser	Thr	Ser	Ser 45	Ser	Ser	Cys	Ile	Ile 50	Ala	Gly	Tyr	Pro	
<b>5</b>		GAG	GAC	እ ጥር	double		7.00	<i>a</i>	C3 III	<b>a</b> a		~~~				55
	245															CAC
10	GIU	GIU	qeA	Met	Pne 60	ALA	Ser	GIN	His	His 65	Arg	Gly	His	His	His 70	His
10	CAC	CAC	CAC	CAT	CAC	CAC	CAT	CAG	CAG	CAG	CAG	CAC	CAG	GCT	CTG	CAA
	273		His													
15				75					80					85	204	<b>4.11</b>
	ACC 341	AAC	TGG	CAC	CTC	CCG	CAG	ATG	TCT	TCC	CCA	CCG	AGT	GCG	GCI	CGG
	Thr	Asn	Trp 90	His	Leu	Pro	Gln	Met 95	Ser	Ser	Pro	Pro	Ser 100	Ala	Ala	Arg
20	CAT	AGC	CTC	TGC	CTC	CAG	CCC	GAC	TCT	GGA	GGG	CCC		GAG	. மாட	GGG
grade Section	389		Leu													
		105		•			110			1	<b>4-1</b>	115		GIU.	Leu	GIĂ
<b>25</b>	AGC 437	AGC	CCG	CCC	GTC	CTG	TGC	TCC	AAC	TCT	TCC	AGC	TTG	GGC	TCC	AGC
<b>30</b>	Ser 120	Ser	Pro	Pro	Val	Leu 125	Сув	Ser	Asn	Ser	Ser 130	Ser	Leu	Gly	Ser	Ser 135
Section 1	ACC 485	CCG	ACT	GGG	GCC	GCG	TGC	GCG	CCG	GGG	GAC	TAC	GGC	CGC	CAG	GCA
		Pro	Thr	Gly	Ala	Ala	Cys	Ala	Pro		Asp .	Tyr	Gly	Arg	Gln	Ala
35	ama	mas.	ccm	000	140					145	_				150	
dining the state of the state o	533															AGC
35 35 40	Leu	ser	Pro	155	GLu	Ala	Glu	Lys	Arg 160	Ser	Gly	Gly	Lys	Arg 165	Lys	Ser
#* **	GAC 581	AGC	TCA	GAC	TCC	CAG	GAA	GGA	AAT	TAC	AAG	TCA	GAA	GTC	AAC	AGC
45	Asp	Ser	Ser 170	Asp	Ser	Gln	Glu	Gly 175	Asn	Tyr	Lys	Ser	Glu 180	Val	Asn	Ser .
10	AAA 629	ccc	AGG	AAA	GAA	AGG	ACA	GCA	TTT	ACC	AAA	GAG	CAA	ATC	AGA	. GAA
	Lys	Pro 185	Arg	Lys	Glu	Arg	Thr 190	Ala	Phe	Thr	Lys		Gln	Ile	Arg	Glu
50·	CTT		GCA	GAA	ىلىنلىن	GCC		CAT	<u>አ</u> አጥ	መአጠ	CTC	195	303	ama	1.00	CGA
	0/1															
<b>55</b>	Leu 200	GLU	nra	GIU.	rne	205	птя	ura	ASI	TYT	210	Thr	Arg	Leu	Arg	Arg 215
	125															GTC
	Tyr	Glu	Ile	Ala	Val 220	Asn	Leu	Asp	Leu	Thr 225	Glu	Arg	Gln	Val	Lys 230	Val
60	TGG	TTC	CAA	AAC	AGG	CGG	ATG	AAG	TGG		AGG	Gта	AAC	GGm		CAG
	773		Gln													
65				235	9	9		-1 -	240	-y s	urd	AGT	-1 A R	245	GTĀ	GTII
×	CAA 821	GGA	GCT	GCG	GCT	CGG	GAA	AAG	GAA	CTG	GTG	AAT	GTG	AAA	AAG	GGA

22311/00114	2	2.	3	1	1	/	0.	0	1	1	4
-------------	---	----	---	---	---	---	----	---	---	---	---

		,	<b>.</b>						22							
	Gln	Gly	Ala 250	Ala	Ala	Arg	Glu	Lys 255	Glu	Leu	Val	Asn	Val 260	Lys	Lys	Gly
5	ACA	CTT	CTC	CCA	TCA	GAG	CTG	TCG	GGA	ATT	GGT	GCA	GCC	ACC	CTC	CAG
<b>9</b>	869 Thr	Leu 265	Leu	Pro	Ser	Glu	Leu 270	Ser	Gly	Ile	Gly	Ala 275	Ala	Thr	Leu	Gln
10	CAA	ACA	GGG	GAC	TCT	ATA	GCA	AAT	GAA	GAC	AGT	CAC	GAC	AGT	GAC	CAC
10	917 Gln 280	Thr	Gly	Asp	Ser	Ile 285	Ala	Asn	Glu	Asp	Ser 290	His	Asp	Ser	Asp	His 295
		TCA	GAG	CAC	GCC	CAC	CTC	TGA								
15	941 Ser	Ser	Glu	His	Ala 300	His	Leu									
20	(2)	INFO	ORMA:	rion	FOR	SEQ	ID 1	۱O:4:	:	-						
25		,	(i) - £	(A)	LEN	GTH:	: 302 amino	ERIST 2 ami 5 aci Linea	.no a		3					
		( :	Li) 1	MOLE	ULE	TYPE	E: pr	otei	.n							
30		. (3	ki) S	SEQUE	ENCE	DESC	RIPI	CION:	SEÇ	Z ID	NO:4	<b>.</b> :				
	Met 1	Glu	His	Pro	Leu 5	Phe	Gly	Сув	Leu	Arg 10	Ser	Pro	His	Ala	Thr 15	Ala
35	Gln	Gly	Leu	His 20	Pro	Phe	Ser	Gln	Ser 25	Ser	Leu	Ala	Leu	His 30	Gly	Arg
The state of the s	Ser	Asp	His 35	Met	Ser	Tyr	Pro	Glu 40	Leu	Ser	Thr	Ser	Ser 45	Ser	Ser	Cys
40	Ile	Ile 50	Ala	Gly	Tyr	Pro	Asn 55	Glu	Glu	Asp	Met	Phe 60	Ala	Ser	Gln	His
45	His 65	Arg	Gly	His	His	His 70	His	His	His	His	His 75	His	His	His	Gln	Gln 80
	Gln	Gln	His	Gln	Ala 85	Leu	Gln	Thr	Asn	Trp 90	His	Leu	Pro	Gln	Met 95	Ser
50 [°]	Ser	Pro	Pro	Ser 100	Ala	Ala	Arg	His	Ser 105	Leu	Сув	Leu	Gln	Pro 110	Asp	Ser
4	Gly	Gly	Pro 115	Pro	Glu	Leu	Gly	Ser 120	Ser	Pro	Pro	Val	Leu 125	Сув	Ser	Asn
55	Ser	Ser 130	Ser	Leu	Gly	Ser	Ser 135	Thr	Pro	Thr	Gly	Ala 140	Ala	СЛа	Ala	Pro
60	Gly 145	Asp	Tyr	Gly	Arg	Gln 150	Ala	Leu	Ser	Pro	Ala 155	Glu	Ala	Glu	Lys	Arg 160
	Ser	Gly	Gly	ГЛЗ	Arg 165	Lys	Ser	Asp	Ser	Ser 170	Asp	Ser	Gln	Glu	Gly 175	Asn
65	Tyr	Lys	Ser	Glu 180	Val	Asn	Ser	Lys	Pro 185	Arg	ГЛя	Glu	Arg	Thr 190	Ala	Phe

Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His Asn

			195					200					205			
5	Tyr	Leu 210	Thr	Arg	Leu	Arg	Arg 215	Tyr	Glu	Ile	Ala	Val 220	Asn	Leu	Asp	Leu
F	Thr. 225	Glu	Arg	Gln	Val	Lys 230	Val	Trp	Phe	Gln	Asn 235	Arg	Arg	Met	Lys	Trp 240
ro	Lys	Arg	Val	Lys	Gly 245	Gly	Gln	Gln	Gly	Ala 250	Ala	Ala	Arg	Glu	Lys 255	Glu
	Leu v	Val	Asn	Val 260	Lys	Lys	Gly	Thr	Leu 265	Leu	Pro	Ser	Glu	Leu 270	Ser	Gly
L5	Île	Gly	Ala 275	Ala	Thr	Leu	Gln	Gln 280	Thr	Gly	Asp	Ser	Ile 285	Ala	Asn	Glu
	Asp	Ser 290	His	Asp	Ser	Asp	His 295	Ser	Ser	Glu	His	Ala 300	His	Leu		